WHAT IS CLAIMED IS:

1.   A speech processing system, comprising:

an acoustic model;

a composite language model including a rules-
        based model portion and a statistical
        model portion; and

a decoder coupled to the acoustic model and the
        composite language model and configured to
        map portions of a natural language speech
        input to pre-terminals and slots, derived
        from a schema, based on the acoustic model
        and the composite language model.

2.   The speech processing system of claim 1 wherein
the decoder is configured to map portions of the
natural language speech input to the slots based on
the rules-based model portion of the composite
language model.

3.   The speech processing system of claim 1 wherein
the decoder is configured to map portions of the
natural language speech input to the pre-terminals
based on the statistical model portion of the
composite language model.

4.   The speech processing system of claim 1 wherein
the statistical model portion of the composite
language model comprises:

a plurality of statistical n-gram models, one
statistical n-gram model corresponding to
each pre-terminal.

5. The speech processing system of claim 4 wherein
the composite language model supports a vocabulary of
words and wherein the statistical n-gram models are
trained based on training data, and wherein words in
the vocabulary that are not used to train a specific
statistical n-gram model comprise unseen words for
the specific statistical n-gram model.

6. The speech processing system of claim 5 wherein
the statistical model portion of the composite
language model further comprises:
a backoff model portion which, when accessed, is
configured to assign a backoff score to a
word in the vocabulary.

7. The speech processing system of claim 6 wherein
each statistical n-gram model includes a reference to
the backoff model portion for all unseen words.

8. The speech processing system of claim 7 wherein
the backoff model portion comprises:
a uniform distribution n-gram that assigns a
uniform score to every word in the
vocabulary.

9.   The speech processing system of claim 1 wherein the rules-based model portion comprises:

a context free grammar (CFG).

10.   A method of assigning probabilities to word hypotheses during speech processing, comprising:

receiving a word hypothesis;

accessing a composite language model having a

plurality of statistical models and a

plurality of rules-based models;

assigning an n-gram probability, with an n-gram

model, to the word hypothesis if the word

hypothesis corresponds to a word seen

during training of the n-gram model; and

referring to a separate backoff model for the

word hypothesis if the word hypothesis

corresponds to a word unseen during

training of the n-gram model; and

assigning a backoff probability to each word

hypothesis, that corresponds to an unseen

word, with the backoff model.

11.   The method of claim 10 and further comprising:

mapping the word hypotheses to slots derived

from an input schema based on the rules-

based models in the composite language

model.

12.   The method of claim 11 and further comprising:

mapping the word hypotheses to pre-terminals
derived from the input schema based on
probabilities assigned by the n-gram models
and the backoff model in the composite
language model.

13. The method of claim 12 wherein referring to a
separate backoff model comprises:
referring to a uniform distribution n-gram.

14. The method of claim 13 wherein assigning a
backoff probability comprises:
assigning a uniform distribution score to every
word in the vocabulary.

15. A composite language model for use in a speech
recognition system, comprising:
a rules-based model portion accessed to map
portions of an input speech signal to slots
derived from a schema; and
a statistical model portion accessed to map
portions of the input speech signal to pre-
terminals derived from the schema.

16. The composite language model of claim 15 wherein
the statistical model portion comprises:
a plurality of statistical n-gram models, one
statistical n-gram model corresponding to
each pre-terminal.

17. The composite language model of claim 15 wherein the rules-based model portion comprises:

  a context free grammar (CFG).

18. The composite language model of claim 16 wherein the composite language model supports a vocabulary of words and wherein the statistical n-gram models are trained based on training data, and wherein words in the vocabulary that are not used to train a specific statistical n-gram model comprise unseen words for the specific statistical n-gram model.

19. The composite language model of claim 18 wherein the statistical model portion of the composite language model further comprises:

  a backoff model portion which, when accessed, is
    configured to assign a backoff score to a
    word in the vocabulary.

20. The composite language model of claim 19 wherein each statistical n-gram model includes a reference to the backoff model portion for all unseen words.

21. The composite language model of claim 20 wherein the backoff model portion comprises:

  a uniform distribution n-gram that assigns a
    uniform score to every word in the
    vocabulary.